Preferential and Absolute Interactions of Solvent Components with Proteins in Mixed Solvent Systems

HIDEO INOUE, *Shionogi Research Laboratory, Shionogi and Company, Ltd., Fukushima-ku, Osaka, Japan,* and SERGE N. TIMASHEFF, *Pioneering Research Laboratory, Eastern Marketing and Nutrition Research Division, Agricultural Research Service, U.S. Department of Agriculture,* and *Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02154,*
Publication No. 818

Synopsis

An equation is derived relating total and preferential interactions of solvent components with macromolecules in a three-component system. Application of this equation to literature data shows that binding of non-polar solvents to proteins parallels the unfolding of the latter. This observation is discussed in terms of local inter-residue and residue–solvent interactions.

Studies on the relation between conformational changes in proteins caused by organic helix-inducing denaturants and the interactions between proteins and the solvent components in water–organic solvent mixtures have indicated that the two phenomena are closely related.[1,2] A direct quantitative comparison between the degree of conformational transition and absolute protein solvent component interactions is complicated, however, by the dependence of the reference state on solvent composition. This follows from the fact that thermodynamic techniques used to measure interactions in three-component systems result only in preferential interactions, i.e., in measurements of the change in chemical potential of the added solvent component, due to its interaction with protein, relative to its chemical potential in the bulk mixed solvent, in the absence of protein.[2,3] In order to understand the relations between protein–solvent interactions and protein conformation in solution, it is necessary to know the absolute extents of interaction between the protein and each solvent component in a solvent mixture. In this paper, an equation will be derived which relates preferential and absolute interactions of solvent components with proteins and it will be applied to data available in the literature.

Using the notation of Scatchard[4] and Stockmayer,[5] let us designate water, the principal solvent, as component 1, the macromolecular solute as component 2, and the added organic solvent as component 3. All concentra-

tions will be expressed in units of grams of component i per gram of water. The thermodynamic measurements of interactions, whether done by light scattering,[6] partial specific volumes,[7] isopiestic equilibrium,[8] or other techniques, give as a final result the change in the activity coefficient of component 3 due to the addition of component 2, without specifying the mechanism of the interactions on a molecular level. The observed interactions represent, therefore, a summation over an entire spectrum of degrees of attraction and repulsion between solvent molecules and regions of the protein chain. Thus, there may be solvent molecules totally immobilized at discrete regions of the protein, with their degrees of rotational and translational motions greatly restricted; other molecules may be partly restricted by weaker interactions which affect their freedom of motion only momentarily; other molecules may be excluded from the domain of the protein. Since the experiments result only in an average value of the interactions for all molecules of a given solvent component, we shall adopt a working model in which solvent molecules will be regarded as belonging to one of two classes, bound (in the present context, the term "binding" is used in its most general sense, i.e., it indicates only the existence of a net attraction between the solvent molecules and the protein) and free. This in no way restricts the general significance of the derived equation.

Let us introduce one gram of component 1 and g_3 grams of component 3 into a bag formed of a membrane impermeable to the macromolecule, and let us immerse this bag into an infinite volume of a mixed solvent of identical composition. Let us now add inside the membrane an infinitesimal amount, dg_2' grams, of component 2. If A_1 is the number of grams of component 1 which becomes bound per gram of protein and A_3 is the number of grams of component 3 which becomes bound per gram of protein, then the amounts of solvent components in the free state inside the bag decrease by A_1dg_2' and A_3dg_2'. The concentration of free component 3 changes from g_3 to g_3' where

$$g_3' = (g_3 - A_3dg_2')/(1 - A_1dg_2') \tag{1}$$

At chemical equilibrium between the inside and the outside solvents at constant temperature, T, and pressure, p, the chemical potentials of component 3 inside, $\mu_3^{(i)}$, and outside, $\mu_3^{(o)}$, the bag must be equal

$$\mu_3^{(i)} = \mu_3^{(o)} = RT \log g_3^{(i)} + RT \log \gamma_3^{(i)} = RT \log g_3^{(o)} + RT \log \gamma_3^{(o)} \tag{2}$$

If it is assumed that the entire difference between the activity coefficients of component 3 inside and outside the bag is due to interactions with protein inside the bag, we may set $\gamma_3^{(i)} = \gamma_3^{(o)} f_3^{(i)}$, where $f_3^{(i)}$ is the fraction of component 3 inside the bag which is not bound to protein. It follows then, that $g_3^{(i)} f_3^{(i)} = g_3^{(o)}$, i.e., at equilibrium, the concentration of unbound component 3 inside the bag must be equal to the concentration of that component outside the bag. In order to maintain this equality, $dg_3 = (g_3 -$

g_3') grams of component 3 per gram of free water must diffuse into the bag through the membrane. Using equation 1, we obtain

$$dg_3 = g_3 - g_3' = (A_3 - g_3A_1)dg_2'/(1 - A_1dg_2') \quad (3)$$

The concentration of component 2 inside the bag at equilibrium, is

$$dg_2 = dg_2'/(1 - A_1dg_2') \quad (4)$$

The preferential interaction parameter, $(\partial g_3/\partial g_2)_{T,p,\mu_3}$, is then

$$(\partial g_3/\partial g_2)_{T,p,\mu_3} = A_3 - g_3A_1 \quad (5)$$

If the absolute degree of hydration of the protein, A_1, is known, equation 5 makes it possible to calculate A_3 the absolute amount of organic solvent bound to the protein, since

$$A_3 = (\partial g_3/\partial g_2)_{T,p,\mu_3} + g_3A_1 \quad (6)$$

This thermodynamically rigorous equation is essentially identical with the one previously deduced from qualitative considerations[1,2] and may serve to prove its validity. If the extent of hydration, A_1, of the protein is known, equation 6 may be used to calculate the absolute amount of organic solvent bound to a protein. Such calculations were carried out for several proteins using available preferential interaction data[1,2,6] and the results were compared with those on conformational changes induced by the same solvents, as a function of solvent composition. In these calculations, it was assumed that the polar groups which were hydrated in the native protein in water remained hydrated to the same extent when the protein unfolded. In addition, an allowance was made for the hydration of newly exposed polar groups as the protein unfolded. The increase in degree of polar group exposure was assumed to parallel the conformational transition, monitored by optical rotatory dispersion and circular dichroism.

The "degree of hydration" of a protein is essentially a vague concept; it is a function of the technique used, since hydrodynamic,[9] x-ray,[10] NMR,[11] and other methods each measure different effects of the macromolecules on the water molecules. For the purpose of our calculations, we have chosen for A_1 the results of water vapor binding at isopiestic equilibrium, reported by Bull and Breese,[12] since these are close to the averages of values given by a variety of techniques. For the native proteins used in the present studies, the hydrations, expressed as grams of water bound per gram of protein, are: 0.32 for β-lactoglobulin A (β-Lg A), 0.32 for bovine serum albumin (BSA), 0.25 for lysozyme, and 0.24 for insulin. Since the dielectric constants of the pure organic solvents used in this study are of the order of 30, in mixtures with water they should exceed the value below which ion pair formation takes place[13] and neutralization of ionized groups occurs. Therefore, the ability of ionized residues to immobilize water molecules should not decrease appreciably with an increase in organic solvent concentration within the range used in our studies. On the other hand, unfolding of the protein should unmask some originally solvent

inaccessible polar groups. Assuming that the net contribution to hydration of the newly exposed groups at the completion of unfolding is of the order of 15 to 20%, it was possible to estimate A_1 at various solvent compositions and using this value to calculate A_3, as well as the total amount of solvent adsorption, $A = A_1 + A_3$, and the composition of the solvation layer, A_3/A_1. The results of these calculations are presented in Tables I and II.

The variation in A_3 as a function of solvent composition is compared with the change in apparent helix contents in β-Lg A in Figure 1 for the system water–2-chloroethanol. It is evident that, except for a short region at low organic solvent contents, the conformational change parallels almost exactly the absolute degree of interaction of the organic solvent with the protein. β-Lg A undergoes a sharp conformational change between 10 and 20% of 2-chloroethanol, i.e., the solvent composition range over which the bulk of

TABLE I
Protein–Solvent Interactions in the Water–2-Chloroethanol System

Chloroethanol, vol. %	g_3, g/g	$\left(\frac{\partial g_3}{\partial g_2}\right)^{\circ}_{T,\mu_1,\mu_3}$	Helix contents,[a] %	A_1, g/g	A_3, g/g	A, g/g	$\frac{A_3}{A_1}$
			β-Lactoglobulin A				
0	0.0	0.0	11	0.32	0.0	0.32	0.0
5	0.063	0.114	10	0.32	0.134	0.454	0.42
10	0.133	0.161	10	0.32	0.204	0.524	0.64
20	0.297	0.466	36	0.36	0.573	0.933	1.59
30	0.505	0.706	46	0.37	0.893	1.263	2.41
40	0.783	0.714	50	0.37	1.004	1.374	2.71
50	1.170	0.637	51	0.37	1.070	1.440	2.89
60	1.748	0.247	53	0.37	0.894	1.264	2.42
80	4.641	−0.624	58	0.37	1.093	1.463	2.96
			Bovine Serum Albumin				
0	0.0	0.0	31	0.32	0.0	0.32	0.0
20	0.297	0.433	51	0.32	0.528	0.848	1.65
40	0.783	0.619	60	0.32	0.870	1.190	2.72
60	1.748	−0.113	58	0.32	0.446	0.766	1.39
80	4.641	−0.787	60	0.32	0.698	1.018	2.18
			Insulin				
0	0.0	0.0	35	0.24	0.0	0.24	0.0
20	0.297	0.359	34	0.24	0.430	0.670	1.79
40	0.783	0.664	35	0.24	0.852	1.092	3.55
60	1.748	0.253	37	0.24	0.672	0.912	2.80
80	4.641	−0.424	37	0.24	0.690	0.930	2.87
			Lysozyme				
20	0.297	0.347		0.25	0.421	0.671	1.68
40	0.783	0.431		0.25	0.627	0.877	2.51
60	1.748	−0.028		0.25	0.409	0.659	1.64
80	4.641	−0.485		0.25	0.675	0.925	2.70

[a]Estimated from b_o of Moffitt-Yang parameters, with $b_o = -630$ for 100% helix and 0 for unordered and β conformations.

TABLE II
Interactions of β-Lactoglobulin with Ethylene Glycol and Methoxyethanol in Aqueous Medium

Organic solvent, vol. %	g_3, g/g	$\left(\frac{\partial g_3}{\partial g_2}\right)^{\circ}_{T,\mu_1,\mu_3}$	Helix contents,[a] %	A_1, g/g	A_3, g/g	A, g/g	$\frac{A_3}{A_1}$
			β-Lg A in H_2O-HOEtOH				
20	0.277	0.013	13	0.32	0.102	0.422	0.32
40	0.731	0.035	8	0.32	0.269	0.589	0.84
60	1.625	0.140	10	0.32	0.660	0.980	2.06
80	4.285	0.419	45	0.37	2.004	2.374	5.42
			β-Lg A in H_2O–MeOEtOH				
20	0.237	0.013	22	0.32	0.089	0.409	0.28
40	0.617	0.000	32	0.34	0.210	0.550	0.62
60	1.351	0.265	58	0.37	0.765	1.135	2.07
80	3.487	0.438	69	0.37	1.728	2.098	4.67

[a]Estimated from b_0 of Moffitt-Yang parameters.

the organic solvent becomes "bound." At higher solvent compositions, both quantities level off, or increase only very slowly. It is quite striking that the binding of organic solvent to the protein occurs already below 10% 2-chloroethanol, i.e., in the regions in which the conformation remains unchanged.

This "induction zone" may be explained in terms of the concept of protein "breathing," [14] i.e., in terms of the thermal fluctuations of the protein conformation.[15] As the protein "breathes," residues, which normally are

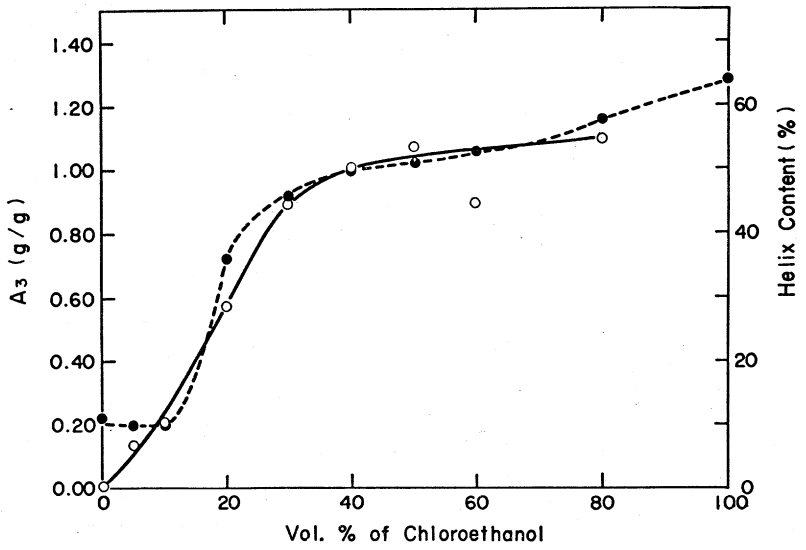

Fig. 1. Variation with solvent composition of solvent binding to β-lactoglobulin A and of apparent helix contents in the water–2-chloroethanol system. Open circles and solid line: binding of 2-chloroethanol to the protein, A_3; filled circles and dashed line: apparent helix contents.

not in contact with solvent, become momentarily exposed to it. Such exposure should permit the non-polar solvent molecules to enter into hydrophobic interactions with non-polar amino acid residues, in particular near the surface of the molecule. Replacement of the internal, inter-residue hydrophobic interactions by external ones should gradually decrease the net stabilization free energy of the protein molecule, until a point at which it is no longer sufficient to balance the destabilizing drive of the conformational entropy. At this point, the molecule should "explode," as is evidenced by a sharp cooperative conformational transition. Such an "explosion," exposing large numbers of previously buried non-polar residues to solvent, should be accompanied by a sharp increase in binding of the non-polar solvent component to the protein, as is indeed observed.

In the case of lysozyme, this gradual exposure of non-polar residues to denaturing solvent, followed by an explosion of the secondary structure, is evidenced by the disruption first of the circular dichroism bands corresponding to interactions between aromatic residues, followed by a general conformational change.[2] This interaction of the non-polar solvent component molecules with the non-polar residues should satisfy the hydrophobic pressure exerted on the latter by the unfavorable water entropy, without the need of folding into an entropically unfavorable globular structure.

As shown in Table I, the amount of 2-chloroethanol bound to protein reaches a value of one gram per gram of β-Lg, i.e., 224 molecules of the non-polar solvent per protein molecule, or approximately 2–3 molecules of the alcohol per non-polar residue. In this system, the sharp conformational transition sets in at a value of A_3 which is at 20% of saturation, i.e., at a point which corresponds to the average solvation by the non-polar solvent of ca. 20 non-polar residues. The net structure stabilizing free energy of a globular protein is of the order of 10–20 kcal/mole[16,17] and the strength of the average inter-residue hydrophobic interactions is 1–2 kcal/mole per residue.[16,18]

These numbers are fully consistent with our model of protein unfolding by non-polar solvents[2]—namely with the mechanism involving the replacement of internal inter-residue hydrophobic interactions by external ones with solvent molecules. Furthermore, these results argue strongly in favor of direct short-range protein–solvent interactions as being a primary cause on the molecular level of protein unfolding, rather than the long-range general effect of the organic solvent on the structure of water in the bulk solvent. The other two solvents used with β-Lg are much weaker denaturing agents, and for them, as seen from Table II, variations in A_3 and apparent α-helix contents increase in essentially parallel fashion, indicating a mechanism of interaction similar to that of 2-chloroethanol.

In the case of the other protein–solvent systems, the information is much less detailed and, therefore, it cannot be subjected to the same type of precise analysis. Nevertheless, it is worth mentioning that, in 2-chloroethanol, the maximal extent of non-polar solvent binding is reached in all

cases at close to 20% alcohol; this is also the solvent composition at which the apparent degree of helicity sharply increases in lysozyme[2] and BSA, as shown in Table I. Insulin appears, at first, to be an anomalous case. Using the optical rotatory dispersion parameter, b_o, as criterion, it would appear that its conformation remains essentially unchanged when transferred from an aqueous medium to chloroethanol[19,20] and yet it binds the organic solvent as strongly as the other proteins.[1] A detailed analysis of the variation of its circular dichroism spectrum with solvent composition,[21] however, has shown that major changes occur in individual bands. The changes, however, are of opposite sign and their summation leads to the observed invariance in b_o. Furthermore, insulin has 3 disulfide bridges for a total of 51 amino acids. The structural constraints imposed by these bridges might be reflected in high-energy barriers to the refolding of large portions of the polypeptide chains into an ordered conformation, such as an α-helix. Thus, while the structure-inducing denaturant might penetrate within the protein molecule, with concomitant conformational alterations and a general loosening of structure, it would not be able to induce the formation of new α-helical regions.

This work was supported in part by the National Institutes of Health Grant No. GM 14603 and the National Science Foundation Grant No. GB 12619.

References

1. S. N. Timasheff and H. Inoue, *Biochemistry*, **7**, 2501 (1968).
2. S. N. Timasheff, *Accts. Chem. Res.*, **3**, 62 (1970).
3. E. F. Casassa and H. Eisenberg, *Advan. Protein Chem.*, **19**, 287 (1964).
4. G. Scatchard, *J. Am. Chem. Soc.*, **68**, 2315 (1946).
5. W. H. Stockmayer, *J. Chem. Phys.*, **18**, 58 (1950).
6. H. Inoue and S. N. Timasheff, *J. Am. Chem. Soc.*, **90**, 1890 (1968).
7. E. F. Woods, S. Himmelfarb, and W. F. Harrington, *J. Biol. Chem.*, **238**, 2374 (1963).
8. E. P. K. Hade and C. Tanford, *J. Am. Chem. Soc.*, **89**, 5034 (1967).
9. C. Tanford, *Physical Chemistry of Macromolecules*, Wiley, New York, 1961, pp. 359, 395.
10. S. N. Timasheff in *Electromagnetic Scattering*, M. Kerker, Ed., Pergamon Press, New York, 1963, p. 337.
11. I. D. Kuntz, Jr., T. S. Brassfield, G. D. Low, and G. V. Purcell, *Science*, **163**, 1329 (1969).
12. H. B. Bull and K. Breese, *Arch. Biochem. Biophys.*, **128**, 488 (1968).
13. S. J. Singer, *Advan. Protein Chem.*, **17**, 1 (1962).
14. A. Hvidt and S. O. Nillsen, *Advan. Protein Chem.*, **21**, 288 (1966).
15. K. V. Linderstrøm-Lang and J. A. Schellman, in *The Enzymes*, P. D. Boyer, H. Lardy, and K. Myrbäck, Eds., Academic Press, New York, 1959, Chap. 10.
16. C. Tanford, *J. Am. Chem. Soc.*, **84**, 4240 (1962).
17. R. Lumry and R. Biltonen, in *Structure and Stability of Biological Macromolecules*, S. N. Timasheff and G. D. Fasman, Eds., Dekker, New York, 1969, Chap. 2.
18. G. Nemethy and H. A. Scheraga, *J. Phys. Chem.*, **66**, 1773 (1962).
19. P. Doty, *Rev. Mod. Phys.*, **31**, 107 (1969).
20. T. T. Herskovits, *J. Biol. Chem.*, **240**, 628 (1965).
21. M. J. Ettinger and S. N. Timasheff, *Biochemistry*, **10**, 831 (1971).